INCREASING THE POWER OF A TWO-LEVEL CHI-SQUARE PROCEDURE OF MULTIPLE HYPOTHESIS TESTING

A.L. Kostevich, I.S. Nikitina
Belarussian State University
Minsk, Belarus
e-mail: kostevich@bsu.by

Abstract

In the paper we consider a two-level chi-square procedure and propose a way to enlarge the power of the procedure under common alternatives. There are given theoretical results for local alternatives and results of the Monte-Carlo experiments.

1 Introduction

When performing statistical simulation, studying random number generators [4], analyzing cryptographic algorithms [5], investigating genetic data [1] researcher often needs to apply a number of statistical tests to a large amount of data. Many statistical procedures are known for multiple hypothesis testing [2].

One-level procedures test the null hypothesis H_0 by applying a number of statistical tests to the same sample. Two-level procedures repeat a first-level test n times over disjoint parts of the sample and obtain a sample of n p-values $P_1, \ldots, P_n$. It is known that under the null hypothesis H_0 the p-values have the uniform distribution $U[0;1]$. Therefore the second level of performing the two-level procedure is to compare an empirical distribution of the p-values $P_1, \ldots, P_n$ with the uniform distribution $U[0;1]$ via a goodness-of-fit test. It should be noted that one-level procedures do not test the local behaviour of random number generators as well as two-level procedures.

A two-level chi-square procedure based on a chi-square goodness-of-fit test is widely used in practice [4, 5]. Having no assumptions about alternatives a basic chi-square procedure divides an interval $[0;1]$ into cells of equal length, calculates the number of p-values in each cell, and calculates the chi-square statistics. This construction does not take into account a fact that under an alternative hypothesis H_1 p-values of many first-level tests are close to 0.

We generalize common deviations from the null hypothesis and propose a mathematical model of the alternative hypothesis. We propose to use unequal cells of partition of the interval $[0;1]$ in order to increase power of the procedure under the specified alternative.

2 Mathematical model

Let $P_1, \ldots, P_n$ be the p-values of the first-level test.

A density of p-values for the first-level test is non-increasing under alternative hypothesis for many testing scenarios [5].

Therefore let us assume that a density of the first-level p-values under the alternative H_1 has the form

$$p_1(x) = wq(x) + (1-w), \quad q(x) = \begin{cases} 1/d_0, & x < d_0 \\ 0, & x \geq d_0 \end{cases}, \tag{1}$$

where d_0 is close to 0. In other words, some p-values are close to 0, and the proportion of these p-values is $w < 1$. Other p-values has the uniform distribution $U[0;1]$.

Now let us consider a well-known chi-square test based on partition of the unit interval into $M+1$ cells:

$$C_{\chi^2}(S) \text{ decides } \begin{cases} H_0, & \text{if } S < \Delta_M(\alpha), \\ H_1, & \text{otherwise}, \end{cases} \qquad S = \sum_{i=0}^{M} \frac{(\nu_i - np_{0,i})^2}{np_{0,i}}, \tag{2}$$

where ν_i is a number of the p-values in the i-th cell of the partition, $p_{0,i}$ is the probability of the event that a p-value lies in the i-th cell, $\Delta_M(\alpha)$ is an α-quantile of the chi-square distribution with M degrees of freedom.

We propose to use the following partition of the unit interval in order to increase power of the procedure under the alternative (1). Let the length of the first interval be d, $d_0 < d < 1/M$, the length of the second interval be $1/M - d$, and the lengths of other $M-1$ intervals be $1/M$. By S_1 and S_2 denote the chi-square statistics based on the specified partition with non-equal cells and the partition with $M+1$ equal cells accordingly:

$$S_1 = \sum_{i=0}^{M} \frac{\left(\nu_i^{(1)} - np_{0,i}^{(1)}\right)^2}{np_{0,i}^{(1)}}, \quad S_2 = \sum_{i=0}^{M} \frac{\left(\nu_i^{(2)} - np_{0,i}^{(2)}\right)^2}{np_{0,i}^{(2)}}.$$

In the next sections we compare the power of chi-square tests $C_{\chi^2}(S_1)$, $C_{\chi^2}(S_2)$ and give results of simulation study.

3 Power comparison for local alternatives

Now consider alternatives that are local:

$$p_1(x) = \frac{w}{\sqrt{n}}q(x) + (1 - \frac{w}{\sqrt{n}}), \quad n \to \infty. \tag{3}$$

Consideration of the sequence of alternatives is induced by the desire to approximate to the power in a region of modest values.

The following lemma gives the statistics distribution under this alternative hypothesis.

Lemma 1. *Suppose the first-level p-values have the density* (3)*; then the statistics S_1 and S_2 have the limiting non-central chi-square distribution with M degrees of freedom and non-centrality parameters $\lambda_1^2 = w^2\frac{1-d}{d}$ and $\lambda_2^2 = w^2 M$ accordingly.*

Proof. The sketch of the proof is following. Elements of the probability vector of $\{\nu_j^{(i)}\}$ are

$$p_{1,j}^{(i)} = p_{0,j}^{(i)} + \frac{\beta_j^{(i)}}{\sqrt{n}},$$

where $\beta^{(1)} = (\beta_0^{(1)}, \ldots, \beta_{M-1}^{(1)})' = (w(1-d), -w(1/M-d), -w/M, \ldots, -w/M)'$, $\beta^{(2)} = (\beta_0^{(2)}, \ldots, \beta_{M-1}^{(2)})' = (w(1-1/(M+1)), -w/(M+1), \ldots, -w/(M+1))'$. It is known that statistics S_1 and S_2 have non-central distribution with a non-centrality parameter

$$\lambda_j^2 = \beta^{(j)'} I(p^{(j)}) \beta^{(j)},$$

where $I(p^{(j)})$ is Fisher-information matrix, $j = 1, 2$. By direct computations we get:

$$I(p^{(1)}) = \left| \begin{array}{ccccc} M + 1/d & M & M & \ldots & M \\ M & M + 1/(1/M - d) & M & \ldots & M \\ M & M & 2M & \ldots & M \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ M & M & M & \ldots & 2M \end{array} \right|$$

and

$$I(p^{(2)}) = \left| \begin{array}{cccc} 2(M+1) & M+1 & \ldots & M+1 \\ M+1 & 2(M+1) & \ldots & M+1 \\ \vdots & \vdots & \ddots & \vdots \\ M+1 & M+1 & \ldots & 2(M+1) \end{array} \right|.$$

Using technique [3] we get the statement of the lemma. □

Now we can compare the power directly. Let $W_1(w)$ be the power of a test with statistic S_1 and $W_2(w)$ be the power of a test with statistic S_2.

Theorem 1. *Under assumptions of Lemma 1 the power of the test with statistic S_1 is more than the power of the test with statistic S_2 for $d < 1/(M+1)$:*

$$W_1(w) > W_2(w), \quad 0 < w < 1.$$

Proof. By lemma 1,

$$W_1(w) = 1 - F_{\chi^2(M;\lambda_1^2)}(\Delta_M),$$
$$W_2(w) = 1 - F_{\chi^2(M;\lambda_2^2)}(\Delta_M).$$

We obviously have $\lambda_1^2 > \lambda_2^2$ for $d < 1/(M+1)$. Therefore, $W_1(w) > W_2(w)$ for $0 < w < 1$. □

4 Simulation study

We compare 3 two-level chi-square procedures based on the statistics S_2 and S_1 with parameter $w = 0.001, 0.01$. The cells count M is 11. The first-level test is the monobit test [5]. For one experiment we generate $n = 580$ 80000-bit samples, n_{alt} samples are generated from the alternative distribution (the probability of 1 is 130/256), other samples are generated from the null distribution. Each sample is tested by the monobit test, producing a p-value. The sample of p-values is tested by the chi-square test with the significance level $\alpha = 0.05$. The power is estimated by 1000 repetitions of the experiment.

In the table below the power of 3 chi-square procedures is given.

n_{alt}	The power		
	$C_{\chi^2}(S_2)$	$C_{\chi^2}(S_1, w = 0.001)$	$C_{\chi^2}(S_1, w = 0.01)$
0	0.043	0.048	0.042
1	0.069	0.069	0.072
3	0.056	0.735	0.102
5	0.06	1	0.218
20	0.349	1	1
50	1	1	1

One can see that the proposed chi-square procedure with non-equal cells is more powerful than the basic procedure.

5 Conclusion

In the paper we propose a way to enlarge the power of the two-level chi-square procedure by appropriate choice of the partition of the interval $[0; 1]$. The obtained results may be generalized for different alternative hypotheses for p-values. The obtained results are substantiated theoretically and illustrated by simulation results.

References

[1] Dudoit S., van der Laan M. J. (2008) *Multiple Testing Procedures with Applications to Genomics*, Springer.

[2] Finner H., Roters M. (2002) Multiple Hypothesis Testing and Expected Number of Type I Errors // *The Annals of Statistics*, Vol. **30**, pp. 220-238.

[3] Kendall M. G., Stuart A. (1973). *The advanced theory of statistics*, vol. II. Griffin, London.

[4] L'Ecuyer P. (1998). *Random Number Generators*. Handbook of Simulation, Wiley.

[5] NIST Special Publication 800-22. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, 2000.